A.BAKER

1632
#3
11-21-00
P.2.

ENTERED

RAW SEQUENCE LISTING DATE: 11/06/2000
PATENT APPLICATION: US/09/639,207 TIME: 11:21:18

Input Set : A:\Sequence Listing - 06618-686001.txt
Output Set: N:\CRF3\11062000\I639207.raw

<110> APPLICANT: Kazemi-Esfarjani, Parsa
Benzer, Seymour
<120> TITLE OF INVENTION: AN ANIMAL MODEL OF POLYGLUTAMINE TOXICITY, METHODS OF USE, AND MODULATORS OF POLYGLUTAMINE TOXICITY
<130> FILE REFERENCE: 06618-686001
<140> CURRENT APPLICATION NUMBER: US 09/639,207
<141> CURRENT FILING DATE: 2000-08-14
<150> PRIOR APPLICATION NUMBER: US 60/148,934
<151> PRIOR FILING DATE: 1999-08-12
<150> PRIOR APPLICATION NUMBER: US 60/148,933
<151> PRIOR FILING DATE: 1999-08-12
<150> PRIOR APPLICATION NUMBER: US 60/177,047
<151> PRIOR FILING DATE: 2000-01-18
<150> PRIOR APPLICATION NUMBER: US 60/205,720
<151> PRIOR FILING DATE: 2000-05-19
<160> NUMBER OF SEQ ID NOS: 69
<170> SOFTWARE: FastSEQ for Windows Version 4.0
<210> SEQ ID NO: 1
<211> LENGTH: 508
<212> TYPE: PRT
<213> ORGANISM: Human
<400> SEQUENCE: 1

```
Met Asp Asp Glu Val Ile Glu Ile Ser Asp Ser Glu Arg Glu Glu Thr
 1               5                  10                  15
Ser Ser Asn Ser Glu Met Asp Val Glu Ile Thr Thr Glu Gln Pro Thr
            20                  25                  30
Ile Asp Val Lys Ala Glu Gln Ile Val Pro Lys Asp Ala Ala Thr Ile
        35                  40                  45
Ala Glu Glu Lys Lys Lys Leu Gly Asn Asp Gln Tyr Lys Ala Gln Asn
    50                  55                  60
Tyr Gln Asn Ala Leu Lys Leu Tyr Thr Asp Ala Ile Ser Leu Cys Pro
65                  70                  75                  80
Asp Ser Ala Ala Tyr Tyr Gly Asn Arg Ala Ala Cys Tyr Met Met Leu
                85                  90                  95
Leu Asn Tyr Asn Ser Ala Leu Thr Asp Ala Arg His Ala Ile Arg Ile
            100                 105                 110
Asp Pro Gly Phe Glu Lys Ala Tyr Val Arg Val Ala Lys Cys Cys Leu
        115                 120                 125
Ala Leu Gly Asp Ile Ile Gly Thr Glu Gln Ala Val Lys Met Val Asn
    130                 135                 140
Glu Leu Asn Ser Leu Ser Thr Ala Val Ala Ala Glu Gln Thr Ala Ala
145                 150                 155                 160
Gln Lys Leu Arg Gln Leu Glu Ala Thr Ile Gln Ala Asn Tyr Asp Thr
                165                 170                 175
Lys Ser Tyr Arg Asn Val Val Phe Tyr Leu Asp Ser Ala Leu Lys Leu
            180                 185                 190
```

RAW SEQUENCE LISTING DATE: 11/06/2000
PATENT APPLICATION: US/09/639,207 TIME: 11:21:18

Input Set : A:\Sequence Listing - 06618-686001.txt
Output Set: N:\CRF3\11062000\I639207.raw

```
 62 Ala Pro Ala Cys Leu Lys Tyr Arg Leu Leu Lys Ala Glu Cys Leu Ala
 63          195                 200                 205
 64 Phe Leu Gly Arg Cys Asp Glu Ala Leu Asp Ile Ala Val Ser Val Met
 65      210                 215                 220
 66 Lys Leu Asp Thr Thr Ser Ala Asp Ala Ile Tyr Val Arg Gly Leu Cys
 67 225                 230                 235                 240
 68 Leu Tyr Tyr Thr Asp Asn Leu Asp Lys Gly Ile Leu His Phe Glu Arg
 69                  245                 250                 255
 70 Ala Leu Thr Leu Asp Pro Asp His Tyr Lys Ser Lys Gln Met Arg Ser
 71              260                 265                 270
 72 Lys Cys Lys Gln Leu Lys Glu Met Lys Glu Asn Gly Asn Met Leu Phe
 73          275                 280                 285
 74 Lys Ser Gly Arg Tyr Arg Glu Ala His Val Ile Tyr Thr Asp Ala Leu
 75      290                 295                 300
 76 Lys Ile Asp Glu His Asn Lys Asp Ile Asn Ser Lys Leu Leu Tyr Asn
 77 305                 310                 315                 320
 78 Arg Ala Leu Val Asn Thr Arg Ile Gly Asn Leu Arg Glu Ala Val Ala
 79                  325                 330                 335
 80 Asp Cys Asn Arg Val Leu Glu Leu Asn Ser Gln Tyr Leu Lys Ala Leu
 81              340                 345                 350
 82 Leu Leu Arg Ala Arg Cys Tyr Asn Asp Leu Glu Lys Phe Glu Glu Ser
 83          355                 360                 365
 84 Val Ala Asp Tyr Glu Thr Ala Leu Gln Leu Glu Lys Thr Pro Glu Ile
 85      370                 375                 380
 86 Lys Arg Met Leu Arg Glu Ala Lys Phe Ala Leu Lys Lys Ser Lys Arg
 87 385                 390                 395                 400
 88 Lys Asp Tyr Tyr Lys Ile Leu Gly Ile Gly Arg Asn Ala Ser Asp Asp
 89                  405                 410                 415
 90 Glu Ile Lys Lys Ala Tyr Arg Lys Lys Ala Leu Val His His Pro Asp
 91              420                 425                 430
 92 Arg His Ala Asn Ser Ser Ala Glu Glu Arg Lys Glu Glu Glu Leu Lys
 93          435                 440                 445
 94 Phe Lys Glu Val Gly Glu Ala Tyr Ala Ile Leu Ser Asp Ala His Lys
 95      450                 455                 460
 96 Lys Ser Arg Tyr Asp Ser Gly Gln Asp Ile Glu Glu Gln Glu Gln Ala
 97 465                 470                 475                 480
 98 Asp Phe Asp Pro Asn Gln Met Phe Arg Thr Phe Phe Gln Phe Asn Gly
 99                  485                 490                 495
100 Gly Gly Arg Asn Asn Ser Ser Phe Asn Phe Glu Phe
101              500                 505
103 <210> SEQ ID NO: 2
104 <211> LENGTH: 2239
105 <212> TYPE: DNA
106 <213> ORGANISM: Drosophila
108 <400> SEQUENCE: 2
109 ggcacgagcc actacttcgc atggcacgct tttttccgtg tgctcggttc gttcggccat      60
110 acaaaacaca aaattcaagt ttaaaaacta aataggcaac taaaagggaa gccgcagcga     120
111 ataaagtgat ttgctgaaag agacgtaaga aagttaatcg catcgaaggc accagaaatc     180
112 ggggatttct aacacggcgc gcgtgcgacg tacatacata cgcaagcgca cacacacacg     240
```

RAW SEQUENCE LISTING DATE: 11/06/2000
PATENT APPLICATION: US/09/639,207 TIME: 11:21:18

Input Set : A:\Sequence Listing - 06618-686001.txt
Output Set: N:\CRF3\11062000\I639207.raw

```
113 aacaattact tgccattgac gcaaaagcga aaaagcagtg gaataaaggg gaattgacaa      300
114 ataacaacgt tttgcaagca ctggactctg gtcgctggtg ttctttcatt ttgtaattgc      360
115 cacgcatgga cgacgaagta attgaaatta gcgacagcga acgcgaagaa acctcatcga      420
116 actccgaaat ggatgtggaa ataacgacag aacagccaac catcgatgtc aaagcagagc      480
117 aaattgtgcc caaggacgcg gcaaccattg ccgaggagaa gaagaaactg ggcaacgacc      540
118 aatacaaggc gcagaactat cagaatgcac tcaagctcta cacggatgcc atatcgctgt      600
119 gtccggactc ggcggcatac tatggcaatc gggccgcctg ctacatgatg ctgctcaact      660
120 ataatagcgc cctgaccgac gcccgacacg ccatacgcat cgatccgggc ttcgagaagg      720
121 cctacgtccg tgtggccaag tgctgtctgg ccctgggcga cattattggc accgaacagg      780
122 ccgtcaaaat ggtcaacgag ctgaattcgc ttagcacggc tgttgctgcc gaacagacgg      840
123 cggcgcaaaa gttgcgccaa ttggaggcca ccattcaggc gaactacgat acgaaatcct      900
124 atcgcaatgt ggtcttctat ttggatagtg ccttgaaatt ggcgcccgcc tgtttgaaat      960
125 atcgtctact caaggctgag tgccttgcat ttttggggcg atgtgatgag gccttggaca     1020
126 ttgcggtcag tgtaatgaaa ctggatacca catcggcgga tgcgatatac gtgagaggtc     1080
127 tgtgcctgta ctacacggac aaccctggaca agggaattct tcatttcgag cgcgccctga     1140
128 ccctcgaccc ggaccactac aagtccaagc agatgcgcag caaatgcaag cagctcaagg     1200
129 agatgaagga gaacggcaat atgctattca agtcgggtcg gtatcgcgag gcacacgtta     1260
130 tctacacgga cgccctgaag atcgatgaac acaacaagga tatcaattcg aaattgcttt     1320
131 acaatcgggc tttggtcaac acgcgtattg gcaatttgcg agaggccgtg gccgattgca     1380
132 atcgagtgct ggagctgaat agtcagtatc tgaaggctct gttgctgcga gcgcgctgct     1440
133 acaatgatct ggagaagttc gaggagtcgg tggcggacta tgagacggcg ctgcagctgg     1500
134 agaagacgcc ggagattaag cgaatgctgc gcgaggccaa gtttgcgttg aagaagtcga     1560
135 agcgaaagga ctactacaag atcctgggca ttggacgcaa tgcgtccgac gacgagatca     1620
136 agaaggcgta tcgcaaaaag gcgctggtac atcatccgga tcgacacgca aacagcagtg     1680
137 ccgaggagcg caaggaggag gagctcaagt tcaaggaggt gggcgaggcg tacgccatac     1740
138 tgtcggatgc tcacaagaag tcgcgctacg acagcggcca ggatatcgag gagcaggagc     1800
139 aagccgactt cgatccgaat caaatgttcc gcacattctt ccaattcaac ggcggtggcc     1860
140 ggaataattc atcgttcaac tttgagttct aggatcccaa cgagtgttgt tcaccaccac     1920
141 agagaagaag accatctcaa tcccatactt tctgcctcat ccgaaaccaa catacagcag     1980
142 cgcacaaatt ttgaactctt ttacatattt cttttccaaa aagcaagaaa ataccacatt     2040
143 ttgattatgt taacgaatga atatatgcca agttatttga aaaaatattc taaatcaaaa     2100
144 taatgcaact aaatttccag tgtaagttca cattttaaa tgttctttct tggatttttt     2160
145 tttcggcaac attaataaat catgggagat ttgtgttaaa taaacagaaa tatacatata     2220
146 aaaaaaaaaa aaaaaaaaa                                                2239
148 <210> SEQ ID NO: 3
149 <211> LENGTH: 273
150 <212> TYPE: PRT
151 <213> ORGANISM: Drosophila
153 <400> SEQUENCE: 3
154 Met Ser Leu Phe Gly Ala Leu Met Gly Asp Phe Asp Asp Asp Leu Gly
155  1               5                  10                  15
156 Leu Met Asn Asn His Met Asn His Thr Met Asn Ala Met Asn Met Gln
157             20                  25                  30
158 Met Arg Ser Met Asn Arg Leu Met Asn Ser Phe Met Pro Asp Pro Phe
159         35                  40                  45
160 Met Gln Val Ser Pro Phe Asp Gln Gly Phe Gln Gln Asn Ala Leu Met
161     50                  55                  60
162 Glu Arg Pro Gln Met Pro Ala Met Pro Ala Met Gly Leu Phe Gly Met
163 65                  70                  75                  80
```

RAW SEQUENCE LISTING DATE: 11/06/2000
PATENT APPLICATION: US/09/639,207 TIME: 11:21:18

Input Set : A:\Sequence Listing - 06618-686001.txt
Output Set: N:\CRF3\11062000\I639207.raw

```
164 Pro Met Met Pro Asn Phe Asn Arg Leu Leu Asn Ala Asp Ile Gly Gly
165                 85                  90                  95
166 Asn Ser Gly Ala Ser Phe Cys Gln Ser Thr Val Met Thr Met Ser Ser
167             100                 105                 110
168 Gly Pro Asp Gly Arg Pro Gln Ile Tyr Gln Ala Ser Thr Ser Thr Lys
169         115                 120                 125
170 Thr Gly Pro Gly Gly Val Arg Glu Thr Arg Arg Thr Val Gln Asp Ser
171     130                 135                 140
172 Arg Thr Gly Val Lys Lys Met Ala Ile Gly His His Ile Gly Glu Arg
173 145                 150                 155                 160
174 Ala His Ile Ile Glu Lys Glu Gln Asp Met Arg Ser Gly Gln Leu Glu
175                 165                 170                 175
176 Glu Arg Gln Glu Phe Ile Asn Leu Glu Glu Gly Glu Ala Glu Gln Phe
177             180                 185                 190
178 Asp Arg Glu Phe Thr Ser Arg Ala Ser Arg Gly Ala Val Gln Ser Arg
179         195                 200                 205
180 His His Ala Gly Gly Met Gln Ala Ile Met Pro Ala Arg Pro Ala Ala
181     210                 215                 220
182 His Thr Ser Thr Leu Thr Ile Glu Pro Val Glu Asp Asp Asp Asp Asp
183 225                 230                 235                 240
184 Asp Asp Asp Cys Val Ile Gln Glu Gln Gln Pro Val Arg Ser Ser Ala
185                 245                 250                 255
186 Gly Arg His Tyr Ser Ser Ala Pro Thr Ala Pro Gln Asn Arg Tyr Asn
187             260                 265                 270
188 Tyr
191 <210> SEQ ID NO: 4
192 <211> LENGTH: 1753
193 <212> TYPE: DNA
194 <213> ORGANISM: Drosophila
196 <400> SEQUENCE: 4
197 ggcacgagga aaatattcgt gaaaattctg catacggaaa gaagaaaatt cgagcaacag      60
198 aaagccaaca caatccacaa aaatgtcttt attcggagcg ttgatgggtg atttcgacga     120
199 cgatctcggc cttatgaaca accacatgaa ccacactatg aacgcgatga acatgcagat     180
200 gcgctcgatg aatcgcctga tgaacagctt tatgcccgat cccttcatgc aggtctcgcc     240
201 ctttgaccag ggattccagc agaacgctct catggagcgt ccgcagatgc cggccatgcc     300
202 agccatgggc ctcttcggca tgcccatgat gccaaacttt aatcgcctgt tgaacgctga     360
203 tattggtggc aattcaggcg catccttctg ccagagcacc gtgatgacca tgtcatcggg     420
204 tcccgatggg cgtcctcaga tctaccaggc cagcactagt accaaaacag gaccgggagg     480
205 cgttcgtgag acccgcagga cggtgcagga ctcgcgcact ggggtgaaga agatggccat     540
206 tggtcatcac atcggcgagc gggcacacat tattgagaaa gagcaggaca tgcgctcagg     600
207 acaactggag gagcgccagg agttcattaa tctggaggag ggagaagccg agcagtttga     660
208 cagggagttt acatcgcgcg ctagtcgcgg agcggtgcag tcaagacatc atgctggtgg     720
209 catgcaggcc atcatgcccg cccgtccagc ggcacacacc tcgacgttga ccattgagcc     780
210 agtggaggac gacgacgacg atgatgatga ctgtgtaatc caggagcagc aaccggttcg     840
211 ctcctccgcg ggccgccatt attccagtgc gccaacggca ccgcagaaca gatataatta     900
212 ctaaatctaa agtcaataca gtatatttta ctaactatcc gataaaacag aaacagaatt     960
213 gcatactata aatttctgct aattacattc ccaactgcgt tcaaacgaaa cgaatatcga    1020
214 atcgaaatca tagaatgcac agagcagcat acatccacat ccctatgccg ccaatccgag    1080
215 gcgccaacaa cgtgccgtaa aacatttcca cacggaggac gaagcggcca gctcctacaa    1140
```

RAW SEQUENCE LISTING DATE: 11/06/2000
PATENT APPLICATION: US/09/639,207 TIME: 11:21:18

Input Set : A:\Sequence Listing - 06618-686001.txt
Output Set: N:\CRF3\11062000\I639207.raw

```
216 ggcggtcaag cgcggcaaga agaagtagta gaaacgtgat catctgtatg ccaacatctt    1200
217 ccgcatcgca cactcaaaaa cactaggaag caaagcgttg ggttctgttc catagcagga    1260
218 aaaccaattc aaatattttt taacaaacac aattctttac cagttctgtc ttatcctgcg    1320
219 tgagtcgacc agaatgcaac actaaaaaat gtacaacttc aagatgctat tgatgtgcac    1380
220 gcaggataca gaacaacttg cttaaattta cttaaaacaa atgtgactat tcaacgccga    1440
221 aatcattaca acacacactc tcagacctaa tcgaaaaatt caatgaaagt aatggaatat    1500
222 atatgaaatc gtaattataa gtttgaatta tttgattaat tctcaagttt ttagattttg    1560
223 ttagccacta agctttaaat tatggatgcc agttagcgtg caaatgaaca caattgattt    1620
224 gaaggctccg aacgatagaa aacaacaatt accaattccc caaatacatg taattcgtaa    1680
225 ggcctaagta aatgttaacg tgaatttaat taaatggtaa ttacattata atagtaaaaa    1740
226 aaaaaaaaaa aaa                                                      1753
228 <210> SEQ ID NO: 5
229 <211> LENGTH: 484
230 <212> TYPE: PRT
231 <213> ORGANISM: Human
233 <400> SEQUENCE: 5
234 Met Ala Ala Thr Glu Pro Glu Leu Leu Asp Asp Gln Glu Ala Lys Arg
235  1               5                  10                  15
236 Glu Ala Glu Thr Phe Lys Glu Gln Gly Asn Ala Tyr Tyr Ala Lys Lys
237             20                  25                  30
238 Asp Tyr Asn Glu Ala Tyr Asn Tyr Tyr Thr Lys Ala Ile Asp Met Cys
239         35                  40                  45
240 Pro Lys Asn Ala Ser Tyr Tyr Gly Asn Arg Ala Ala Thr Leu Met Met
241     50                  55                  60
242 Leu Gly Arg Phe Arg Glu Ala Leu Gly Asp Ala Gln Gln Ser Val Arg
243 65                  70                  75                  80
244 Leu Asp Asp Ser Phe Val Arg Gly His Leu Arg Glu Gly Lys Cys His
245                 85                  90                  95
246 Leu Ser Leu Gly Asn Ala Met Ala Ala Cys Arg Ser Phe Gln Arg Ala
247             100                 105                 110
248 Leu Glu Leu Asp His Lys Asn Ala Gln Ala Gln Gln Glu Phe Lys Asn
249         115                 120                 125
250 Ala Asn Ala Val Met Glu Tyr Glu Lys Ile Ala Glu Thr Asp Phe Glu
251     130                 135                 140
252 Lys Arg Asp Phe Arg Lys Val Val Phe Cys Met Asp Arg Ala Leu Glu
253 145                 150                 155                 160
254 Phe Ala Pro Ala Cys His Arg Phe Lys Ile Leu Lys Ala Glu Cys Leu
255                 165                 170                 175
256 Ala Met Leu Gly Arg Tyr Pro Glu Ala Gln Ser Val Ala Ser Asp Ile
257             180                 185                 190
258 Leu Arg Met Asp Ser Thr Asn Ala Asp Ala Leu Tyr Val Arg Gly Leu
259         195                 200                 205
260 Cys Leu Tyr Tyr Glu Asp Cys Ile Glu Lys Ala Val Gln Phe Phe Val
261     210                 215                 220
262 Gln Ala Leu Arg Met Ala Pro Asp His Glu Lys Ala Cys Ile Ala Cys
263 225                 230                 235                 240
264 Arg Asn Ala Lys Ala Leu Lys Ala Lys Lys Glu Asp Gly Asn Lys Ala
265                 245                 250                 255
266 Phe Lys Glu Gly Asn Tyr Lys Leu Ala Tyr Glu Leu Tyr Thr Glu Ala
```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY DATE: 11/06/2000
PATENT APPLICATION: US/09/639,207 TIME: 11:21:19

Input Set : A:\Sequence Listing - 06618-686001.txt
Output Set: N:\CRF3\11062000\I639207.raw

L:1278 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:1298 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:1338 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19
L:1374 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:1429 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24
L:1430 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24
L:1455 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25
L:1535 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:1540 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:1626 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35
L:1649 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36
L:1686 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:38
L:1761 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42
L:1783 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43
L:1840 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46
L:1912 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51
L:1917 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51
L:1941 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52
L:1961 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:53
L:2029 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56
L:2140 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:62